

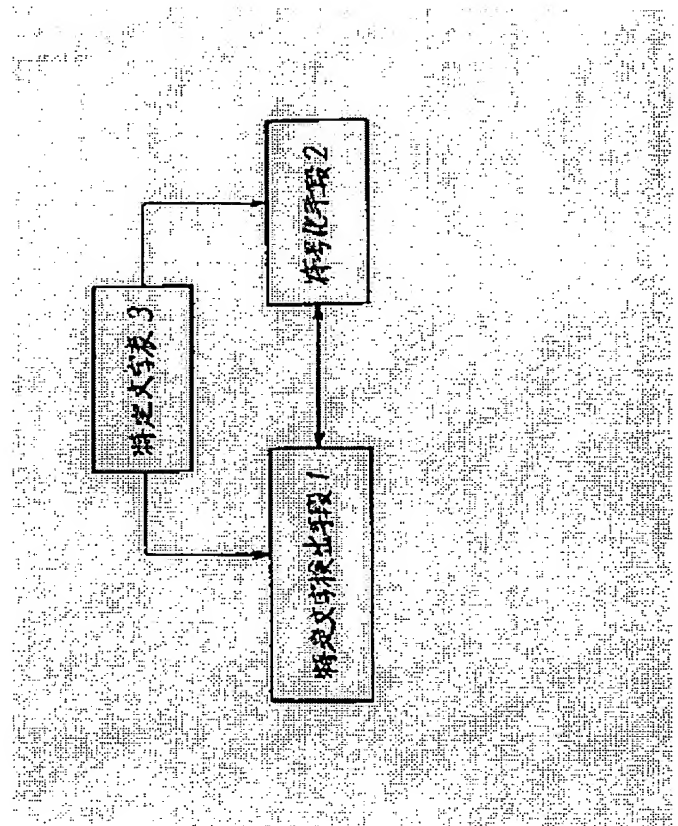
JAPANESE WORD INFORMATION COMPRESSION SYSTEM

Patent number: JP2271723
Publication date: 1990-11-06
Inventor: KOSAKAI TAKAO
Applicant: NEC CORP
Classification:
- international: H03M7/30
- european:
Application number: JP19890093457 19890412
Priority number(s):

Abstract of JP2271723

PURPOSE:To attain a high compression rate by retrieving Japanese word information expressed in a 2-byte code based on a list on which special characters are registered, and allocating a code to a detected Japanese character.

CONSTITUTION:A special character detection means 1 calls a coding means 2 to apply coding when a Japanese character is registered in a specific character list 3. Moreover, the special character detection means 1 stores the Japanese character in the built-in memory as it is when the character is not registered in the list 3. The coding means 2 represents the location of the entry into the character list 3 equal to the present Japanese character in a binary number, sets the most significant bit (0 to 1) and sends the result as a conversion code and sends it to the detection means 1. The detection means 1 stores the received conversion code into the built-in memory. Thus, when Japanese words using lots of kinds such as Hiragana (round form of Japanese syllabary), symbols and numerals are compressed, a high compression rate is obtained.



⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A)

平2-271723

⑤ Int. Cl.⁵

識別記号

庁内整理番号

⑬ 公開 平成2年(1990)11月6日

H 03 M 7/30

6832-5J

審査請求 未請求 請求項の数 1 (全4頁)

⑭ 発明の名称 日本語情報圧縮方式

⑰ 特 願 平1-93457

⑱ 出 願 平1(1989)4月12日

⑲ 発 明 者 小 堺 隆 男 東京都港区芝5丁目33番1号 日本電気株式会社内

⑳ 出 願 人 日本電気株式会社 東京都港区芝5丁目7番1号

㉑ 代 理 人 弁理士 内 原 晋

明 細 書

発明の名称

日本語情報圧縮方式

特許請求の範囲

2バイトのコードで表現された日本語情報を圧縮する日本語情報圧縮方式において、

(A) あらかじめ定められた特定文字が登録された表にもとづき、前記2バイトのコードで表現された日本語情報を検索して前記特定文字の日本語文字を検出する特定文字検出手段、

(B) 前記検出された日本語文字に符号を割り当てる符号化手段、

を備えたことを特徴とする日本語情報圧縮方式。

発明の詳細な説明

(産業上の利用分野)

本発明は日本語情報圧縮方式に関し、特に2バ

イトのコードで表現された日本語情報を圧縮する日本語情報圧縮方式に関する。

(従来の技術)

従来、日本語情報を圧縮する場合、日本語文字が14ビットで表現することができ、2バイトに格納するときにはできる2ビットの余りを詰めて圧縮する方法が通常とられている。

(発明が解決しようとする課題)

上述した従来の日本語情報圧縮方式は、約1.14倍の圧縮効果しか得られないという問題点がある。

本発明の目的は、上記の圧縮率より大きい圧縮率を得ることができる日本語情報圧縮方式を提供することにある。

(課題を解決するための手段)

本発明の日本語情報圧縮方式は、2バイトのコードで表現された日本語情報を圧縮する日本語情報圧縮方式において、

(A) あらかじめ定められた特定文字が登録された表にもとづき、前記2バイトのコードで表現

された日本語情報を検索して前記特定文字の日本語文字を検出する特定文字検出手段、

(B) 前記検出された日本語文字に符号を割り当てる符号化手段、

を備えて構成されている。

〔実施例〕

次に、本発明の実施例について図面を参照して説明する。

第1図は本発明の一実施例のブロック図である。

第1図に示す日本語情報圧縮方式は、特定文字表3を検索して特定のコードの日本語文字を検出する特定文字検出手段1、日本語文字に符号を割り当てる符号化手段2から構成されている。

次に、動作を説明する。

第2図は特定文字検出手段1の動作を示す流れ図である。

第2図のステップ21で処理の対象となる日本語文字の先頭に位置付け、ステップ22で現在の位置の日本語文字が特定文字表3に登録されてい

る文字であれば、ステップ23で符号化手段2を呼び出して符号化を行う。ステップ22で現在の位置の日本語文字が特定文字表3に登録されている文字でなければ、ステップ24でその日本語文字をそのまま内蔵したメモリに格納する(この場合は圧縮が行われない)。ステップ25では処理を終了するか否かの判定が行われ、“NO”であれば、ステップ22に戻り検索を続行する。

次に、第1図の符号化手段2における動作を説明する。

第3図は符号化手段2の動作を示す流れ図である。

第3図のステップ31で現在の日本語文字と等しい特定文字表3のエントリの位置を2進数で表わし、その最上位ビットを「ON」(0→1)にしたものを変換符号として特定文字検出手段1に送出し、特定文字検出手段1は受信した変換符号を内蔵されたメモリに格納する。

第4図は本発明の一実施例の処理過程を示す説明図であり、第5図は本発明の一実施例の特定文

字表3の例を示す説明図である。

第4図においては、日本語文字列として、「“標準”の在り方」をJIS-X-0208“情報交換用漢字符号系”の日本語コード(16進数)で表現し、圧縮を行った例を示している。例示した日本語文字に含まれる文字のうち、第5図の特定文字表3に示す「”」、「”」、「の」及び「り」は、そのエントリの位置を1バイトで表現して最上位ビットを「ON」とした形式で符号化がなされ、圧縮結果の個所で示されるような結果が得られる。この場合は、最初4桁で表現されていた「”」、「”」、「の」及び「り」のコードがそれぞれ2桁のコードに圧縮されている。

〔発明の効果〕

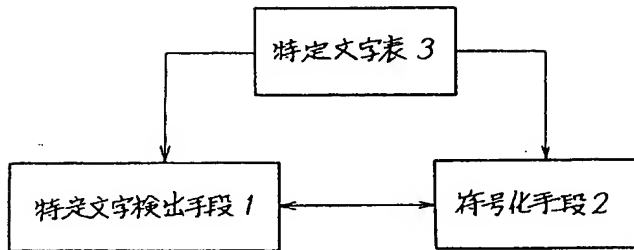
以上説明したように、本発明は、ひらがな、記号及び数字等が比較的多数使用されている日本語文字の圧縮において高い圧縮率を得ることができるという効果を有する。

図面の簡単な説明

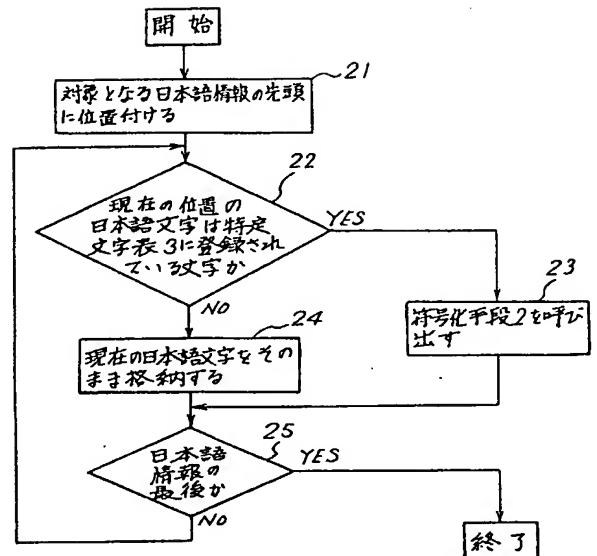
第1図は本発明の一実施例のブロック図、第2図は特定文字検出手段の動作を示す流れ図、第3図は符号化手段の動作を示す流れ図、第4図は本発明の一実施例の処理過程を示す説明図、第5図は本発明の一実施例の特定文字表の例を示す説明図である。

1…特定文字検出手段、2…符号化手段、3…特定文字表。

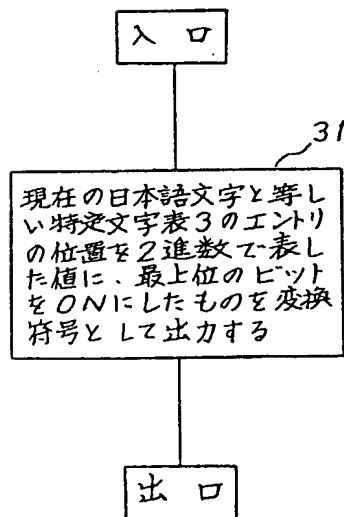
代理人 弁理士 内 原 晋



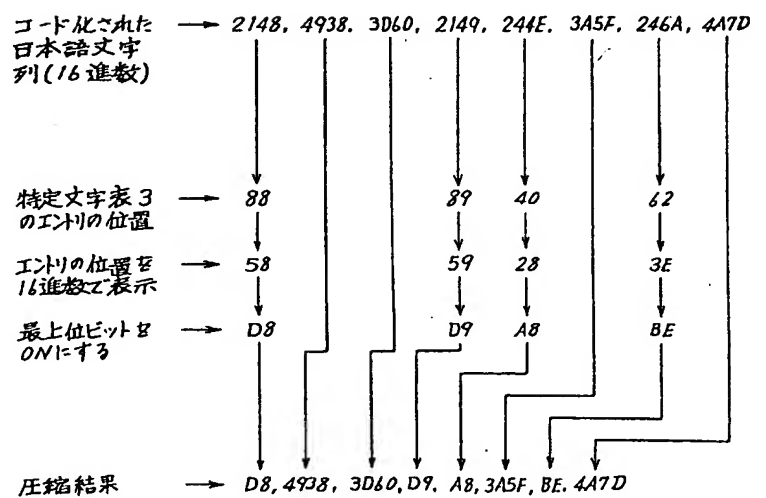
第 1 図



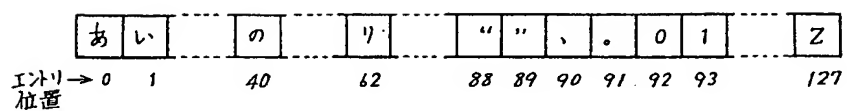
第 2 図



第 3 図



第 4 図



第 5 図